



**HAL**  
open science

## Negative regulation of the Nrf1 transcription factor by its N-terminal domain is independent of Keap1: Nrf1, but not Nrf2, is targeted to the endoplasmic reticulum

Yiguo Zhang, Dorothy H. Crouch, Masayuki Yamamoto, John D. Hayes

### ► To cite this version:

Yiguo Zhang, Dorothy H. Crouch, Masayuki Yamamoto, John D. Hayes. Negative regulation of the Nrf1 transcription factor by its N-terminal domain is independent of Keap1: Nrf1, but not Nrf2, is targeted to the endoplasmic reticulum. *Biochemical Journal*, 2006, 399 (3), pp.373-385. 10.1042/BJ20060725 . hal-00478597

**HAL Id: hal-00478597**

**<https://hal.science/hal-00478597>**

Submitted on 30 Apr 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

17/07/06

Version 3 of BJ 2006/0725

**Negative regulation of the Nrf1 transcription factor by its N-terminal domain is independent of Keap1: Nrf1, but not Nrf2, is targeted to the endoplasmic reticulum**

Yiguo Zhang§¶, Dorothy H. Crouch§, Masayuki Yamamoto#, and John D. Hayes§¶

§*Biomedical Research Centre, Ninewells Hospital and Medical School, University of Dundee, Dundee DD1 9SY, Scotland, United Kingdom, and #Center for Tsukuba Advanced Research Alliance and Institute of Basic Medical Sciences, University of Tsukuba, Tsukuba 305-8577, Japan.*

**Running title: Nrf1 and Nrf2 are regulated by different mechanisms**

**Keywords:** Nrf1, Nrf2, Keap1, Antioxidant response element, Endoplasmic reticulum, Oxidative stress

¶ To whom correspondence should be addressed:

Dr. Yiguo Zhang and Prof. John D. Hayes  
Biomedical Research Center  
Ninewells Hospital and Medical School  
University of Dundee  
Dundee DD1 9SY  
Scotland  
United Kingdom  
Tel: 0044 1382 632788  
Fax: 0044 1382 669993  
Email: [y.z.zhang@dundee.ac.uk](mailto:y.z.zhang@dundee.ac.uk) and [john.hayes@cancer.org.uk](mailto:john.hayes@cancer.org.uk)

## **ABSTRACT**

Nrf1 (Nuclear Factor-Erythroid 2 p45 subunit-related factor 1) and Nrf2 regulate antioxidant response element (ARE)-driven genes. At its N-terminal end, Nrf1 contains 155 additional amino acids (aa) that are absent from Nrf2. This 155-aa polypeptide includes the N-terminal domain (NTD, aa 1-124) and a region (aa 125-155) that is part of acidic domain 1 (aa 125-295). Within acidic domain 1, residues 156-242 share 43% identity with the Neh2 degron of Nrf2 that serves to destabilize this latter transcription factor through an interaction with Keap1 (Kelch-like ECH-associated protein 1). We have examined the function of the 155-aa N-terminal polypeptide in Nrf1 along with its adjacent Neh2-like subdomain. Activation of ARE-driven genes by Nrf1 was negatively controlled by the NTD through its ability to direct Nrf1 to the endoplasmic reticulum. Ectopic expression of wild-type Nrf1 and mutants lacking either the NTD or portions of its Neh2-like subdomain into wild-type and mutant mouse embryonic fibroblasts indicated that Keap1 controls neither the activity of Nrf1 nor its subcellular distribution. Immunocytochemistry showed that whereas Nrf1 gave primarily cytoplasmic staining, that was coincident with that of an endoplasmic reticulum marker, Nrf2 gave primarily nuclear staining. Attachment of the NTD from Nrf1 to the N-terminus of Nrf2 produced a fusion protein that was redirected from the nucleus to the endoplasmic reticulum. Although this NTD/Nrf2 fusion protein exhibited less transactivation activity than wild-type Nrf2, it was nevertheless still negatively regulated by Keap1. Thus, Nrf1 and Nrf2 are targeted to different subcellular compartments and are negatively regulated by distinct mechanisms.

## INTRODUCTION

Cells express many inducible genes that allow adaptation to oxidative stress. These include those encoding NAD(P)H:quinone oxidoreductase 1 (NQO1<sup>1</sup>), glutathione *S*-transferase, the glutamate cysteine ligase modifier (GCLM) subunit, ferritin and heme oxygenase-1 [1-3]. Transcriptional activation of these genes occurs through antioxidant response elements (AREs) present in their promoter regions and is mediated by the cap'n' collar (CNC) basic-region leucine zipper (bZIP) proteins Nrf1 (Nuclear Factor-Erythroid 2 p45 subunit-related factor 1) and Nrf2 [4]. These two factors both bind AREs as heterodimers with small Maf proteins [5, 6]. Together, Nrf1 and Nrf2 regulate ARE-driven genes such as *nqo1* and *gclm* [7, 8].

Nrf1 and Nrf2 make distinct contributions to the antioxidant status of cells *in vivo*. Global knockout (KO) of Nrf1 has been shown to result in death of mice *in utero* [9, 10]. Liver-specific KO of Nrf1 in mice at birth led to non-alcoholic steatohepatitis that developed into hepatoma [11]. Unlike disruption of *nrf1*, disruption of *nrf2* yields mice that develop normally and do not spontaneously develop cancer [12], though they are more sensitive to chemical carcinogenesis [13, 14]. Given the facts that Nrf1 and Nrf2 regulate common genes and they are both expressed in the liver [15], it is surprising that hepatic KO of these factors in mice yields different phenotypes. It therefore appears that *in vivo* Nrf2 is unable to compensate for loss of Nrf1 in mouse liver, but the reason for this is not known.

Whilst it is axiomatic that the distinct biological activities of Nrf1 and Nrf2 are determined by their primary structures, it is unclear what specific features of the bZIP proteins are responsible for them producing markedly different phenotypes in the KO mice. Lack of detailed information about the structure and function of Nrf1 represents an obvious impediment to our understanding of the mechanism(s) by which it makes its unique contribution to maintenance of redox status. In this context, a complicating factor is that Nrf1 exists as multiple isoforms. In the human, a long version of Nrf1 has been described, designated TCF11 (transcription factor 11), which comprises 772 amino acids (aa) [16, 17]. Human Nrf1 protein is slightly shorter than TCF11 and contains 742 aa; it arises by alternative splicing of exon 4 [16-18]. Despite this difference, Nrf1 and TCF11 both transactivate ARE-driven reporter genes to a similar extent [19]. The mouse does not possess an orthologue of TCF11, rather it only expresses a full-length Nrf1 protein of 741 aa that shares 95% sequence identity with human Nrf1 [20].

The most obvious feature that distinguishes Nrf1 from Nrf2, and may therefore account for their distinct biological functions, is the presence of an additional 155 N-terminal aa polypeptide in the former bZIP protein. This N-terminal extension in Nrf1 includes 31 aa that were considered to be part of the N-terminal acidic domain in TCF11 [16, 17]. The function of neither the N-terminal 124 aa nor residues 125-155 in Nrf1 is known.

An additional explanation for the putative distinct functions of Nrf1 and Nrf2 is that they are differentially regulated by Keap1 (Kelch-like ECH-associated protein 1), a substrate adaptor for the Cullin-3:Rbx1 E3 ubiquitin ligase [21, 22]. Keap1 controls the stability of Nrf2 ensuring that the bZIP protein is held at low levels under normal homeostatic conditions but can rapidly accumulate during oxidative stress when the substrate

adaptor function of Keap1 is inhibited [23, 24]. Nrf2 interacts with Keap1 through its N-terminal Neh2 domain [25, 26]. This interaction involves binding through both a high-affinity ETGE motif and a low-affinity DLG motif in the Neh2 domain of Nrf2 [27, 28]. In Nrf1, the regions corresponding to the high- and low-affinity Keap1-binding sites in Nrf2 are conserved. Using a quantitative yeast two-hybrid assay, Kobayashi et al [29] have found that Keap1 can interact with Nrf1, albeit with only about 28% of the affinity Keap1 has for Nrf2. However, following transfection into mouse 3T3 cells, these workers also found that Keap1 did not influence transactivation of a reporter gene by Nrf1 [29]. It is therefore uncertain whether Keap1 regulates Nrf1.

In this paper we have sought to identify: (i) the function of the additional 155 aa N-terminal region in Nrf1 that is absent from Nrf2; (ii) whether Nrf1 is negatively regulated by Keap1 through its Neh2-like region located between aa 156-242.

## EXPERIMENTAL

### Generation of Nrf1 expression constructs

The cDNA fragment coding full-length Nrf1 was obtained using the ProSTAR ultra HF RT-PCR system (Stratagene, La Jolla, CA) supplemented with mouse kidney RNA and primers Nrf1F and Nrf1R. The cDNA product was purified and subcloned into the *KpnI/XbaI* site of pcDNA3.1/V5His B (Invitrogen) that translates into a C-terminally V5-tagged protein. Nrf2 expression constructs have been described previously [23]. The fidelity of the cDNA product and all constructs made in this study were sequence-verified by the Human Genome Group, Department of Molecular and Cellular Pathology, University of Dundee. All oligonucleotide primers (see Supplementary Table 1) were synthesized by MWG Biotech Co (Ebersberg, Germany).

### Mutagenesis of Nrf1 and Keap1 constructs

Two strategies were employed for mutagenesis. Firstly, relatively short internal deletion mutants of Nrf1, encoding Nrf1<sup>Δ125-170</sup>, Nrf1<sup>ΔETGE</sup> and Nrf1<sup>ΔDIDLID/DLG</sup> were produced using the QuickChange Site Directed Mutagenesis kit (Stratagene) as described previously [30]; relevant pairs of sense and antisense primers are listed in Supplementary Table 1. Secondly, the cDNA fragments encoding N- and C-terminally truncated Nrf1 proteins (Nrf1<sup>Δ2-120</sup>, Nrf1<sup>Δ2-150</sup> and Nrf1<sup>Δ2-170</sup>) were generated by PCR through utilizing their corresponding primers (see Supplementary Table 1); these amplified fragments were also subcloned into the *KpnI/XbaI* sites of pcDNA3.1/V5His B.

### Constructs for Nrf1 fused with Gal4 or Nrf2

The pcDNA3.1Gal4D-V5 based expression constructs encoding the Gal4 DNA-binding domain (Gal4D) fused to various Nrf1 mutants were generated by first subcloning the Gal4D-encoding fragment of pGBT9 (Clontech Laboratories, Inc.) into the *HindIII* site of a modified pcDNA3.1/V5His B plasmid. Eleven different length cDNA fragments encoding different portions of Nrf1 were produced using PCR and equivalent GN primer pairs (see Supplementary Table 1) and cloned into the *BamHI/EcoRI* sites of pcDNA3.1Gal4D-V5. The transactivation activity of these Gal4D fusion constructs was determined using the Gal4 luciferase reporter

$P_{TK}$ -UAS-Luc [31]. The Gal4D chimeras containing the NTD were subjected to immunocytochemistry. In order to obtain NTD/Nrf2 chimaeras, the *XhoI/XbaI* fragment of Nrf1/pcDNA3.1/V5His B was replaced by either the PCR product for Nrf2<sup>f</sup> (full-length mouse Nrf2 with Met<sup>1</sup> and Met<sup>2</sup> mutated into Val) or the PCR product for Nrf2<sup>s</sup> (residues 16-597 of Nrf2 with Met<sup>17</sup> mutated into Ile); the resulting pcDNA3.1/V5His B expression constructs encoded chimaeric proteins containing 125 N-terminal aa from Nrf1 fused to either residues 1-597 of Nrf2 (giving NTD/Nrf2<sup>f</sup>) or to residues 16-597 of Nrf2 (giving NTD/Nrf2<sup>s</sup>).

### Reporter constructs

The  $P_{1016/nqo1}$ -Luc and a mutant (mut1) containing a scrambled ARE across nucleotides -450 to -415 within the same construct have been described elsewhere [32]. Three different 41-bp ARE-driven luciferase reporter constructs [i.e.,  $P_{SV40nqo1}$ -ARE-Luc,  $P_{SV40gclm}$ -ARE-Luc, and  $P_{SV40porphobilinogen\ deaminase\ (PBGD)}$ -ARE-Luc] were made by insertion of the appropriately hybridized complementary oligonucleotides [32-34] with 2-bp overhangs into the *XhoI/BglIII* restriction sites of a modified pGL3-Promoter reporter vector. The pGL-6×ARE was a gift from Dr. Jun Wang (unpublished) and was engineered by inserting 6 copies of the oligonucleotide 5'-CCCGTGACAAAGCA-3' into pGL3-Basic (Promega); the core ARE from the promoter of the rat *glutathione S-transferase Alpha2* gene is underlined [35]. The Gal4 luciferase reporter  $P_{TK}$ -UAS-Luc was a gift from Dr. Mattias Gustafsson (unpublished) and was made by cloning the *SalI/XhoI* fragment of pLacZr, containing a module of Gal4-specific binding sites (UAS×4) linked to the minimal thymidine kinase promoter (TK) of herpes simplex virus [31], into pGL3-Basic (Promega) cleaved with *XhoI*. The pcDNA3.1/V5His/*lacZ*, pcDNA4/HisMax/*lacZ* (Invitrogen) and pRL-TK (Promega) were used to control for transfection efficiency. On some occasions a  $P_{TK}$ -ARE-Luc reporter plasmid was used to reduce background transactivation; this construct was made by replacing the *BglIII/HindIII* fragment containing the SV40 promoter in  $P_{SV40nqo1}$ -ARE-Luc reporter with an equivalent region encompassing the TK promoter from the pRL-TK.

### Cell culture, transfection and luciferase reporter assays

Rat liver RL-34 epithelial cells and monkey kidney COS-1 cells were grown as described previously [23, 32]. Mouse embryonic fibroblasts (MEFs) from wild-type (*keap1*<sup>+/+</sup>) and Keap1 KO (*keap1*<sup>-/-</sup>) mice [36] were grown in Iscove's Modified Dulbecco's medium (Invitrogen) supplemented with 10% (v/v) fetal calf serum, 10 µg/ml of insulin, 5.6 µg/ml transferrin, 6.7 ng/ml of sodium selenite, 0.25% (w/v) NaHCO<sub>3</sub>, 100 U/ml of penicillin and 100 µg/ml streptomycin at 37°C in humidified air with 5% CO<sub>2</sub>. Equal numbers of cells (3 × 10<sup>5</sup> cells/well) were seeded and allowed to grow for 18-24 h in 6-well plates. Once cells reached ~60% confluence, they were transfected using Lipofectin reagent or Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Approximately 36-48 h following transfection, gene reporter activity was determined using the Dual Luciferase Reporter Assay System (Promega). The ARE- and UAS-driven firefly luciferase activities were normalized to that of the Renilla reporter, or β-galactosidase (gal) reporter constructs used to control for transfection efficiency [23]. After subtraction of background reporter activity, obtained following cotransfection

with ARE-null reporter (i.e., pGL3-Basic) and empty expression vectors, Nrf1-mediated activation activity was calculated. Significant differences in the transactivation activities were determined using the Student's *t* test.

### Western blotting

Cells were disrupted in 1 × RIPA buffer [50 mM TrisHCl, pH 7.5, 150 mM NaCl, 1% IPEGAL (Nonidet P-40), 0.5% sodium deoxycholate, 0.1% sodium dodecyl-sulphate (SDS), 1 mM dithiothreitol and a protease inhibitor cocktail]. The clarified cell lysates were examined using two electrophoretic systems. In the first method, samples were diluted with 5 × SDS sample buffer (pH 6.8) and then heated at 100°C for 10 min before being subjected to separation by SDS-PAGE in a Tris-glycine discontinuous system [37]. In the second method, cell lysates were diluted in 4 × LDS (lithium dodecyl-sulfate) sample buffer (pH 8.4) and heated at 70°C for 10 min before being resolved by LDS-NuPAGE gel (Invitrogen). Subsequently, western blot analysis was performed as reported previously [30]. C-terminally tagged Nrf1 and Nrf2 proteins were located with primary antibodies against the V5 epitope tag (Invitrogen) and visualized using peroxidase-labelled secondary antibody against mouse IgG. Immunoblotting using a GAPDH antibody (Amino Bioproducts Ltd, Frelstedt, Germany) served as a loading control.

### Immunocytochemistry and confocal microscopy

Cells ( $1 \times 10^5$ ) were grown on a cover-slip in 6-well plates. Approximately 24 h after transfection, they were fixed with 4% paraformaldehyde in 20 mM sodium phosphate buffer (pH 6.8). The cells were washed 4 times, for 3 min, in phosphate buffer before each immunocytochemistry step. Fixed cells were permeabilized by incubation in 0.2% Triton X-100 before non-specific antibody binding sites were blocked with 5% FBS and 0.5% normal goat serum. The cover-slips were incubated with a primary antibody against V5 tag or Nrf1 for 1 h at 4°C before being incubated with green-fluorescent Alexa-Fluor 488 (FITC) for an additional 1 h at 4°C. Nuclear DNA was stained in a solution containing 10 µg/ml of 4', 6-diamidino-2-phenylindole (DAPI). The confocal images were obtained using a LSM 510 laser scanning microscope system (Carl Zeiss, Germany). To address the question of whether Nrf1 can localize to the endoplasmic reticulum, the pDsRed2-ER (BD Biosciences Clontech, CA) vector was employed that encodes a fusion of *Discosoma sp.* red fluorescent protein (DsRed) with the endoplasmic reticulum targeting sequence of calreticulin coupled to its N-terminus and the endoplasmic reticulum retention sequence tetrapeptide KDEL linked to its C-terminus.

## RESULTS

### Identification of domains in Nrf1

Bioinformatic analysis of Nrf1 suggested that it is a modular protein. As shown in Table 1, we have divided it into nine domains. Nrf1 is most obviously distinguished from Nrf2 by the presence of the NTD and an acidic/polar region (residues 125-155) located at its N-terminus. We included the latter region in acidic domain 1 as has been done previously during characterization of TCF11 [16, 17]. Within the acidic domain 1, and

immediately adjacent to the additional 155 aa polypeptide in Nrf1, is a stretch of amino acid residues (156-242), that shares 43% identity with the N-terminal Neh2 region of Nrf2. Most significantly, this Neh2-like (Neh2L) subdomain in Nrf1 contains DLG and ETGE motifs that in the case of Nrf2 are responsible for its interaction with Keap1 [23,27]. This paper is concerned with the function of the NTD and the Neh2L subdomain in Nrf1.

### **Nrf1 is less potent than Nrf2 at transactivating gene expression**

Although both Nrf1 and Nrf2 positively regulate ARE-driven gene expression [4], they may display selectivity in the genes they recognize. The AREs in the mouse *nqo1* and *gclm* gene promoters were chosen for study because examination of *nrf1* and *nrf2* KO MEFs suggests *nqo1* and *gclm* are targets for Nrf1 and Nrf2 [7]. Also, the ARE in human *PBGD* was studied because previously a random oligonucleotide-based screen to select transcription factor binding sites showed its ARE is optimal for Nrf1-MafG binding [38]. Figure 1A shows that Nrf1 activated a *P*<sub>-1016/nqo1</sub>-ARE-Luciferase reporter gene about 2.5-fold, but did not activate the *P*<sub>-1016/nqo1</sub>-mut1 reporter gene. By comparison, a parallel experiment revealed that Nrf2 transcriptionally activated the same mouse *nqo1*-ARE-Luciferase reporter gene about 5.0-fold.

Based on the mass of DNA transfected, Nrf1 exhibited less transactivation activity than Nrf2 towards three reporter genes driven by 41-bp ARE-containing sequences from the promoters of *nqo1*, *gclm* and *PBGD* (Figure 1B). To determine whether differences in protein level might account for the lower potency of Nrf1, immunoblotting experiments were undertaken. Surprisingly, although Nrf1 was less active than Nrf2, it clearly appeared to be the more abundant of the two proteins following transfection into COS-1 cells (Figure 1C). It was also noted that Nrf1 was resolved by SDS/PAGE into two or three major closely migrating polypeptides whereas examination of Nrf2 in the same analysis gave a single western blot band. Presumably, Nrf1 is subject to some type of post-translational modification that does not occur in the case of Nrf2. In addition to the major electrophoretic Nrf1 bands, three relatively minor smaller polypeptides were also observed that, based on published work, probably arise from translation being initiated from internal in-frame ATG codons [17-19, 39].

### **The NTD negatively regulates transactivation by Nrf1**

To explore why Nrf1 is less potent than Nrf2 at transactivating ARE-driven gene expression, various deletion mutants of Nrf1 were generated and tested in transfection experiments. The most striking difference we observed occurred following omission of the NTD from Nrf1. The resulting factor exhibited significantly greater transactivation activity than full-length Nrf1; this effect was more obvious in RL-34 cells than in COS-1 cells. As shown in Figure 2A, removal of residues 2-120 from Nrf1 caused a substantial increase in transactivation by the truncated transcription factor in both cell lines. Further deletion of N-terminal amino acids up to residues 170 caused no additional increase in transactivation activity. It was consistently noted that Nrf1<sup>Δ2-120</sup> affected similar levels of transactivation as Nrf2, on the basis of mass of both expression vector transfected and protein levels expressed in cells (data not shown). This indicates that the NTD is probably the major negative regulating domain in Nrf1.

To further test whether the NTD negatively controls Nrf1 activity, expression constructs were generated for the Gal4D protein fused to either full-length Nrf1 or one of a series of truncated proteins (Figure 2B, left upper panel). Transfection of COS-1 cells with constructs expressing the Gal4D protein fused at its C-terminus to a series of polypeptides generated from Nrf1<sup>1-607</sup>, but lacking increasingly large portions from the N-terminus, showed that removal of residues 1-44 from Nrf1 had a little effect on Gal4/UAS-driven reporter gene expression (Figure 2B, c.f. lanes 2 and 3). However, removal of residues 1-65 or 1-119 caused a gradual step-like increase in transactivation of Gal4 reporter activity from ~600- to 1050-fold. Maximal activation occurred when almost the entire NTD was deleted (Figure 2B, lane 8). No obvious difference in the level of the different ectopic fusion proteins was observed (Figure 2B, lower three panels).

### **The DIDLID/DLG and ETGE sequences in the Neh2L subdomain of Nrf1 neither influence its transactivation activity nor its cytoplasmic localization**

Results presented in Figure 2 show that the NTD negatively regulates Nrf1. However, Nrf1 also contains, in close proximity to the NTD, a region that is similar to the Neh2 degnon in Nrf2; in the case of Nrf2, the Neh2 region controls the half-life of the bZIP protein through an interaction with Keap1 [23]. If Nrf1 were controlled by Keap1, it would be predicted that deletion of the DIDLID/DLG sequence (aa 171-186) or ETGE motif (aa 234-237) in the Neh2L subdomain would increase transactivation by, and nuclear accumulation of, the mutant bZIP protein. To determine whether either these motifs in Neh2L or immediately adjacent N-terminal sequences control the function of Nrf1, we deleted either residues 125-170, the DIDLID/DLG sequence or the ETGE motif from an expression construct encoding the full-length bZIP protein (Figure 3A). Transfection experiments showed that removal of residues 125-170 in Nrf1 significantly decreased reporter gene activity when compared to that obtained by wild-type Nrf1. In the same series of experiments, deletion of the DIDLID/DLG sequence caused a modest decrease in the reporter transactivation activity in COS-1 cells, but this was not observed in RL-34 cells (Figure 3B, upper panels). No significant difference in transactivation of reporter gene activity by Nrf1<sup>ΔETGE</sup> from that by wild-type Nrf1 was observed in either COS-1 or RL-34 cells. As neither an increase in reporter gene activity nor a change in protein levels was observed in these Nrf1 mutants (Figure 3B), we conclude that neither the DIDLID/DLG sequence nor the ETGE motif negatively regulate the transcription factor. The inference from this finding is that it is unlikely Keap1 negatively controls Nrf1.

Confocal microscopy was performed to determine whether the DIDLID/DLG sequence and ETGE motif in Neh2L influence the subcellular localization of Nrf1. In COS-1 cells, ectopic wild-type Nrf1 gave stronger cytoplasmic staining than nuclear staining, whereas ectopic Nrf2 was located primarily in the nucleus (Figure 3C). Quantification of this staining pattern indicated that 70% of wild-type Nrf1 was located in the cytoplasm (Figure 3D). No differences were observed between the subcellular distributions of wild-type Nrf1 and its mutants Nrf1<sup>ΔDIDLID/DLG</sup> and Nrf1<sup>ΔETGE</sup> (Figures 3C and D). Identical results were obtained in RL-34 cells transfected with the same constructs (data not shown). These results indicate that the cytoplasmic-nuclear distribution of Nrf1 is unaffected by deletion of the DIDLID/DLG sequence and ETGE motif. Furthermore, the

finding that Nrf1 is less potent than Nrf2 at transactivating ARE-driven reporter gene expression may be explained by the observation that it is primarily located in the cytoplasm.

### Keap1 does not antagonize Nrf1

In the case of Nrf2, the difference between transactivation of an ARE-reporter gene by the wild-type protein and the  $\Delta$ ETGE mutant is only obvious when the appropriate expression plasmids for the bZIP factor are cotransfected with an expression construct for Keap1 [23]. To test whether this also holds true for Nrf1, we cotransfected COS-1 cells with an ARE-driven reporter gene along with expression constructs for Nrf1 and wild-type Keap1 (Figure 4A). Transactivation of ARE-driven luciferase activity by ectopic Nrf1 was unaffected by cotransfection with Keap1 (Figure 4A, left panel). By contrast, Keap1 reduced transactivation of reporter gene activity by Nrf2 to approximately 50% of that observed by Nrf2 in the absence of ectopic Keap1. Further transfection experiments revealed that activation of a pTK-ARE-Luc reporter gene by ectopic Nrf1 was essentially identical in *keap1*<sup>+/+</sup> and *keap1*<sup>-/-</sup> MEFs (Figure 4A, middle panel); the small increase of luciferase activity in the knockout MEFs may be due to the enhanced endogenous Nrf2 activity in these cells. By contrast, transfection of an expression construct for Nrf2 gave significantly higher reporter gene activity in *keap1*<sup>-/-</sup> MEFs than in *keap1*<sup>+/+</sup> MEFs. We therefore conclude that full-length Nrf1 is not regulated by Keap1.

To test the hypothesis that the Neh2L subdomain is capable of interacting with Keap1 but is prevented from doing so by the presence of the NTD, we transfected MEF cell lines with expression constructs for wild-type Nrf1 and Nrf1 <sup>$\Delta$ 2-120</sup> along with a luciferase reporter plasmid. Consistent with previous data (Figure 2A), Nrf1 <sup>$\Delta$ 2-120</sup> exhibited greater transactivation activity than wild-type Nrf1 (Figure 4A, middle panel). Most importantly, no difference was observed in the transactivation of reporter gene activity by Nrf1 <sup>$\Delta$ 2-120</sup> in *keap1*<sup>+/+</sup> and *keap1*<sup>-/-</sup> MEFs (Figure 4A, middle panel). These data again suggest that Keap1 does not regulate Nrf1, even when the bZIP factor lacked its NTD. This conclusion was supported by the further observation that activation of pGL-6 $\times$ ARE-Luc activity by wild-type Nrf1 or Nrf1 <sup>$\Delta$ 2-120</sup> was not significantly affected by the reintroduction of Keap1 into *keap1*<sup>-/-</sup> MEFs in a rescue type experiment (Figure 4A, right panel).

Ectopically expressed Nrf1 showed a similar sub-cellular distribution in both *keap1*<sup>+/+</sup> and *keap1*<sup>-/-</sup> MEFs (Figure 4B). The predominantly cytoplasmic distribution of wild-type Nrf1 did not differ from that of Nrf1 <sup>$\Delta$ DIDLID/DLG</sup> and Nrf1 <sup>$\Delta$ ETGE</sup> in *keap1*<sup>-/-</sup> MEFs.

### The NTD is essential for cytoplasmic localization of Nrf1

Confocal microscopy was used to determine whether the NTD could control the extra-nuclear localization of Nrf1. Following transfection of mutant forms of Nrf1 into COS-1 and RL-34 cells, it was found that cytoplasmic location of Nrf1 was abolished upon removal of its NTD. The Nrf1 <sup>$\Delta$ 2-120</sup> mutant protein was located primarily in the nucleus rather than cytoplasm (Figure 5A). Deletion of other single domains gave staining that was indistinguishable from wild-type Nrf1 (data not shown). These findings suggest that the NTD targets Nrf1 to a region within the cytoplasm.

In transfection experiments, extra-nuclear ectopic Nrf1 co-localized with the ER/DsRed marker (Figure 5A and B). This finding suggests that full-length Nrf1 is specifically targeted to this organelle. Overall, our data demonstrate that Nrf1 activity and its cytoplasmic localization are negatively regulated by the NTD, and that the negative effect of this domain on transactivation is independent of Keap1.

### The NTD directs Nrf2 fusion proteins to the endoplasmic reticulum

To test whether the NTD is capable of targeting proteins to the endoplasmic reticulum, we generated expression constructs for proteins in which NTD was attached to the N-terminus of wild-type Nrf2. Confocal microscopy showed that in COS-1 and RL-34 cells, ectopically expressed wild-type Nrf2 was located primarily in the nucleus (Figures 3C and 5B). Ectopic wild-type Nrf2 was also located in the nucleus of *keap1*<sup>-/-</sup> MEFs (Figure 6A). Conversely, a predominantly cytoplasmic localization was found following transfection of *keap1*<sup>-/-</sup> MEFs with expression constructs encoding fusion proteins in which the NTD from Nrf1 (residues 1-125) was attached to the N-terminus of either full-length Nrf2 (Nrf2<sup>f</sup>, residues 1-597) or short Nrf2 (Nrf2<sup>s</sup>, residues 16-597). The cytoplasmic staining of NTD/Nrf2<sup>f</sup> and NTD/Nrf2<sup>s</sup> was identical to that observed with ER/DsRed.

Luciferase reporter assays showed that by comparison with wild-type Nrf2, NTD/Nrf2<sup>f</sup> and NTD/Nrf2<sup>s</sup> were less potent at transactivating an ARE-driven reporter gene (Figure 6B). Interestingly, a higher level of transactivation by the NTD/Nrf2 fusion proteins was observed in *keap1*<sup>-/-</sup> MEFs than in wild-type MEFs (Figure 6B), but this increase in activity was prevented by transfection into the mutant MEFs of an expression construct for Keap1 (Figure 6C). These results suggest that the Nrf2 chimaeras are negatively regulated by the NTD through a mechanism that involves tethering to the endoplasmic reticulum. Significantly, these NTD/Nrf2 fusion proteins were found to be regulated by Keap1, suggesting that the presence of the NTD does not obscure the effects of the Neh2 domain within the chimaera.

To determine whether the NTD can redirect the subcellular localization of not only Nrf2 but also an unrelated protein such as Gal4D, we transfected COS-1 cells with the Gal4D/Nrf1 fusion constructs. As shown in Figure 2B, the fusion proteins from Gal4D/Nrf1<sup>1-607</sup> and Gal4D/Nrf1<sup>1-298</sup>, that contained the NTD, were almost exclusively located in the endoplasmic reticulum; conversely the NTD-deficient fusion proteins from Gal4D/Nrf1<sup>120-607</sup> and Gal4D/Nrf1<sup>120-298</sup> were primarily expressed in the nucleus (Figure 6D); they have significant transactivation activity when compared with Gal4D/Nrf1<sup>1-607</sup>. In addition, the Gal4D polypeptide from pcDNA3.1Gal4D-V5 was found to reside exclusively in the nucleus. No change in protein levels expressed by these constructs was observed (data not shown). These data indicate that an endoplasmic reticulum targeting signal exists with the NTD of Nrf1.

## DISCUSSION

Gene KO experiments in the mouse have shown that Nrf1, but not Nrf2, is essential for embryonic development [7, 9, 10, 40]. Furthermore, liver-specific KO of Nrf1 leads to hepatoma, whereas no such phenotype arises in *nrf2*<sup>-/-</sup> mice. The basis for the phenotypic differences in these mutant mice is not known. In an attempt to address this issue we have examined the function of the NTD in Nrf1 because it is absent from Nrf2. In

addition, we have examined whether the Neh2L region in Nrf1 results in it being inhibited by Keap1 because the Neh2 domain in Nrf2 is responsible for this latter bZIP protein being degraded rapidly in a Keap1-dependent fashion under normal homeostatic conditions.

### **Targeting of Nrf1 to the endoplasmic reticulum through its NTD**

Our reporter gene experiments showed that deletion of the entire NTD from Nrf1 caused a substantial increase in ARE-driven transcription, but deletion of other domains in the bZIP protein did not result in similar activation. This indicates that the NTD is the principal negative mode by which the factor is regulated. We have also found that the negative regulation of Nrf1 is associated with an ability of NTD to direct the protein to the endoplasmic reticulum. Thus, whereas ectopic wild-type Nrf1 gave broad cytoplasmic staining when examined by confocal microscopy, Nrf1<sup>Δ2-120</sup> gave almost exclusively nuclear staining when similarly studied. The conclusion that the NTD negatively controls Nrf1 by directing it to the endoplasmic reticulum is supported by our finding that NTD/Nrf2<sup>f</sup>, NTD/Nrf2<sup>s</sup> and NTD/Gal4D chimaeric proteins were all directed to this organelle whereas wild-type Nrf2 and Gal4D were essentially exclusively nuclear proteins. Furthermore, NTD/Nrf2<sup>f</sup> and NTD/Nrf2<sup>s</sup> exhibited less transactivation activity towards the ARE-reporter gene than wild-type Nrf2 suggesting that, as expected, targeting these bZIP proteins to membranes inhibits their activity.

Bioinformatic analyses suggested that the N-terminal part of the NTD is responsible for Nrf1 being directed to the endoplasmic reticulum. Proteins, such as calreticulin, calnexin and calmeglin, which are known to be located in the endoplasmic reticulum, each contain at their N-terminus a region with significant homology to Nrf1. In particular, the peptide sequence GLLQFTILLSLIGVRVD found between residues 11-27 of Nrf1 shares 35% identity with the MLLSVPLLLGLLGLAAA leader sequence found between residues 1-17 of mouse calreticulin [41]. The latter sequence has been implicated in the translocation of calreticulin into the endoplasmic reticulum, suggesting that the GLLQFTILLSLIGVRVD peptide in Nrf1 is probably responsible for it being targeted to the same organelle.

Besides Nrf1, a small number of other transcription factors have been shown to be targeted to the endoplasmic reticulum. These include sterol regulatory element binding proteins SREBP1 and SREBP2 [42], as well as ATF6 and the related cyclic AMP response element-binding proteins CREB4 and CREBH [43-45]. Transcription factors that are located in the endoplasmic reticulum all possess at least one transmembrane domain that anchors them to the organelle in such a fashion that their DNA-binding and transactivation domains are orientated towards the cytosolic face of the membrane. This topology is necessary to ensure that when such factors are released from intracellular membranes they are free to translocate to the nucleus. Release of SREBP and CREBH from the endoplasmic reticulum entails trafficking to the Golgi apparatus and regulated intramembrane proteolysis, a process that involves cleavage of the factors in response to specific stimuli by the sequential actions of Site-1 and Site-2 proteases. It is not known whether Nrf1 is processed by these proteases, but bioinformatic analysis predicts that residues 81-90 form an amphipathic  $\alpha$ -helix and within this region there exists the tetrapeptide RRLL that could be a recognition sequence for the Site-1 peptidase [46-48]. In addition, based on comparison with proteins containing tripartite signal sequences, it is predicted that flanking the

hydrophobic sequence between residues 11-24 in Nrf1 there exists a signal peptidase cleavage site at the C-terminal side of Thr<sup>30</sup> [49]. The presence of proteolytic cleavage sites in the NTD requires further examination, but their existence seems highly probable given the heterogeneity in electrophoretic mobility of Nrf1 observed in Figure 1.

### Conserved motifs in the NTD of Nrf1 and Nrf3

An unexpected finding during this study was that the NTD in Nrf1 appears to be represented in Nrf3. Comparison between the N-terminal regions of Nrf1 and Nrf3 revealed two conserved regions, GLLQxTILLSLxGxRVDxD (where x indicates any amino acid) and RLLxxVRALxxxxPxTxVxAWLVH, found between residues 11-29 and 82-106, respectively, in Nrf1, and residues 12-30 and 76-100 in Nrf3. We refer to these as the N-terminal homology box 1 (NHB1) and NHB2, with the former residing closest to the N-termini of the factors. As noted above, NHB1 is predicted to be part of an endoplasmic reticulum targeting sequence based on its homology with the N-termini of calreticulin, and NHB2 is predicted to form part of an amphipathic  $\alpha$ -helix that may be cleaved by Site-1 protease.

The presence of NHB1 and NHB2 in Nrf3 suggests that it may also be associated with the endoplasmic reticulum. This hypothesis requires to be tested. The remaining two Nrf family members, Nrf2 and NF-E2 p45, appear unlikely to locate to the endoplasmic reticulum because they do not possess sequences related to either NHB1 or NHB2. In fact, as has been demonstrated in Figures 3, 5 and 6 of this paper, Nrf2 resides primarily in the nucleus. Similarly, NF-E2 p45 also locates primarily to the nucleus [50].

### Keap1 does not regulate Nrf1

The Nrf2 transcription factor is negatively regulated by Keap1 through at least two interactions that involve on the one hand the low-affinity DLG motif and the high-affinity ETGE motif in its Neh2 domain, and on the other hand the Kelch-repeat domain in Keap1 [23, 27, 28, 51-54]. These interactions are redox-dependent and result in the stability of Nrf2 being controlled by Keap1 under normal homeostatic conditions [3, 26, 53]. In the Neh2L subdomain of Nrf1, the DIDLID element, the DLG motif and the ETGE motif are all conserved [52]. Prior to the present study it was unclear from the literature whether this bZIP protein is regulated by Keap1; a previous two-hybrid analysis showed that Nrf1 can interact with Keap1 in yeast cells, but that Keap1 does not inhibit Nrf1 transactivation activity in mouse 3T3 cells [29]. Our data show that deletion of the DLG and ETGE motifs in Nrf1 does not influence either its transactivation activity or its cytoplasmic localization. We have found that in wild-type and mutant MEFs, Keap1 does not influence the subcellular distribution of Nrf1 nor does it alter the ability of the bZIP protein to activate ARE-driven gene expression.

In view of the fact that Nrf1 is subject to N-terminal proteolysis, we investigated the possibility that the NTD might somehow mask Keap1-binding sites in the Neh2L subdomain; we postulated that proteolytic removal of the NTD at the endoplasmic reticulum might alleviate some sort of steric hindrance thereby allowing the Neh2L subdomain to interact with Keap1. However, we obtained no experimental data to support this hypothesis. Upon transfection into *keap1*<sup>+/+</sup> and *keap1*<sup>-/-</sup> MEFs, we found that transactivation of ARE-driven

gene expression by Nrf1<sup>A2-120</sup> was not influenced by the presence of Keap1. During these experiments we did however show that Keap1 can negatively regulate NTD/Nrf2 chimaeric proteins. In this case, attachment of either the NTD to Nrf2<sup>f</sup> (i.e., full-length Nrf2) or the N-terminal 125 aa of Nrf1 to Nrf2<sup>s</sup> (i.e., residues 16-597 of Nrf2) resulted in decreased transactivation activity but it did not obscure the inhibitory effect of Keap1 on transactivation of an ARE-driven luciferase reporter gene. These data indicate that differences between the effect of Keap1 on Nrf1 and Nrf2 cannot be attributed to the presence of the NTD.

Our experiments involving transfection into *keap1*<sup>-/-</sup> MEFs of the Nrf1 NTD deletion mutants show that the presence of Neh2L in Nrf1 is insufficient to result in its inhibition by Keap1. The three following reasons may account for failure of Keap1 to inhibit Nrf1: differences around the ETGE motif in Nrf1 may reduce its affinity for Keap1; differences around the DLG motif in Nrf1 may compromise the interaction with Keap1; loss of critical ubiquitin acceptor lysine residues between the DLG and ETGE motifs in Nrf1 may block ubiquitylation. Given the report that in yeast two-hybrid experiments Nrf1 possesses only about 28% of the affinity for Keap1 than Nrf2 exhibits for Keap1, it appears most likely that changes around the ETGE motif in the Neh2L subdomain of Nrf1 are responsible for the apparent reduced interaction between Nrf1 and Keap1. It has recently been reported that the LDEETGEFL nonapeptide between residues 76-84 of Nrf2 is involved in the high-affinity interaction of the bZIP protein with Keap1 [55]. As this sequence is represented by VDGETGESF (residues 231-239) in Nrf1, it is clear that there are four amino acid changes around the ETGE motif (i.e. Val<sup>231</sup>, Gly<sup>233</sup>, Ser<sup>238</sup>, Phe<sup>239</sup>) that could be responsible for the inability of Keap1 to inhibit the bZIP factor. It is not known which, if any, of these four residues might be responsible for the reduced affinity between Nrf1 and Keap1. It is also possible that changes around the DLG motif in Nrf1 have inactivated the low-affinity Keap1-binding site. In this case, the functional sequence in Nrf2 is QDIDLGVSR (residues 26-34) whereas that in Nrf1 is QDIDLGAGR (residues 181-189); thus the presence of Ala<sup>187</sup> and/or Gly<sup>188</sup> in Nrf1 may be responsible for the loss of Keap1 inhibition. Lastly, it could be argued that the reason Keap1 does not inhibit Nrf1 is because the lysine residues to which ubiquitin is conjugated in Nrf2 are not conserved in Nrf1. In the case of Nrf2, the lysine residues that are targeted for ubiquitin conjugation by cullin3:Rbx1 (i.e. Lys<sup>44</sup>, Lys<sup>50</sup>, Lys<sup>52</sup>, Lys<sup>53</sup>, Lys<sup>56</sup>, Lys<sup>64</sup> and Lys<sup>68</sup>) are located in an  $\alpha$ -helical region, residues 39-71, held between the DLG and ETGE motifs [22, 56, 57]. Sequence alignment shows that amongst the seven lysines in Nrf2 that can accept ubiquitin only two are conserved in Nrf1 (i.e. Lys<sup>199</sup> and Lys<sup>205</sup>). It is not known whether this is insufficient to allow ubiquitylation of Nrf1. It is also possible that the protein fold within the Neh2L region prevents Lys<sup>199</sup> and Lys<sup>205</sup> from being available for ubiquitylation. These possibilities should be explored in future.

### Concluding comments

In this paper we have discovered that transactivation by Nrf1 is negatively regulated by its NTD through a mechanism that entails an association of the factor with the endoplasmic reticulum. We have provided evidence that the negative regulation of Nrf1-mediated transactivation and control of its cytoplasmic location are independent of Keap1. Our findings that Nrf1 can associate with the endoplasmic reticulum and is not regulated by Keap1, whereas Nrf2 does not associate with this organelle and is regulated by Keap1, suggest that these two

bZIP factors mediate adaptation to redox stress in different subcellular compartments. Presumably this will explain, at least in part, the distinct phenotypes observed in the Nrf1 and Nrf2 KO mice.

## ACKNOWLEDGEMENTS

This work was supported by both the Association for International Cancer Research (grant 03-074) and Tenovus Scotland (grant T05/6). We thank Drs. M. McMahon, P. Nioi, L. G. Higgins and T.W.P. Devling for providing expression constructs and for critical advice. We are indebted to Drs. M. Gustafsson and J. Wang for gifts of reporter constructs.

## REFERENCES

- Hayes, J. D., Ellis, E. M., Neal, G. E., Harrison, D. J. and Manson, M. M. (1999) Cellular response to cancer chemopreventive agents: contribution of the antioxidant responsive element to the adaptive response to oxidative and chemical stress. *Biochem. Soc. Symp.* **64**, 141-168
- Nguyen, T., Sherratt, P. J. and Pickett, C. B. (2003) Regulatory mechanisms controlling gene expression mediated by the antioxidant response element. *Annu. Rev. Pharmacol. Toxicol.* **43**, 233-260
- Nguyen, T., Sherratt, P. J., Nioi, P., Yang, C. S. and Pickett, C. B. (2005) Nrf2 controls constitutive and inducible expression of ARE-driven genes through a dynamic pathway involving nucleocytoplasmic shuttling by Keap1. *J. Biol. Chem.* **280**, 32485-32492
- Venugopal, R. and Jaiswal, A. K. (1996) Nrf1 and Nrf2 positively and c-Fos and Fra1 negatively regulate the human antioxidant response element-mediated expression of NAD(P)H:quinone oxidoreductase1 gene. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14960-14965
- Motohashi, H., O'Connor, T., Katsuoka, F., Engel, J. D. and Yamamoto, M. (2002) Integration and diversity of the regulatory network composed of Maf and CNC families of transcription factors. *Gene* **294**, 1-12
- Venugopal, R. and Jaiswal, A. K. (1998) Nrf2 and Nrf1 in association with Jun proteins regulate antioxidant response element-mediated expression and coordinated induction of genes encoding detoxifying enzymes. *Oncogene* **17**, 3145-3156
- Leung, L., Kwong, M., Hou, S., Lee, C. and Chan, J. Y. (2003) Deficiency of the Nrf1 and Nrf2 transcription factors results in early embryonic lethality and severe oxidative stress. *J. Biol. Chem.* **278**, 48021-48029
- Chen, L., Kwong, M., Lu, R., Ginzinger, D., Lee, C., Leung, L. and Chan, J. Y. (2003) Nrf1 is critical for redox balance and survival of liver cells during development. *Mol. Cell. Biol.* **23**, 4673-4686
- Farmer, S. C., Sun, C. W., Winnier, G. E., Hogan, B. L. and Townes, T. M. (1997) The bZIP transcription factor LCR-F1 is essential for mesoderm formation in mouse development. *Genes Dev.* **11**, 786-798
- Chan, J. Y., Kwong, M., Lu, R., Chang, J., Wang, B., Yen, T. S. and Kan, Y. W. (1998) Targeted disruption of the ubiquitous CNC-bZIP transcription factor, Nrf-1, results in anemia and embryonic lethality in mice. *EMBO J.* **17**, 1779-1787
- Xu, Z., Chen, L., Leung, L., Yen, T. S., Lee, C. and Chan, J. Y. (2005) Liver-specific inactivation of the Nrf1 gene in adult mouse leads to nonalcoholic steatohepatitis and hepatic neoplasia. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 4120-4125
- Chan, K., Lu, R., Chang, J. C. and Kan, Y. W. (1996) NRF2, a member of the NFE2 family of transcription factors, is not essential for murine erythropoiesis, growth, and development. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13943-13948
- Ramos-Gomez, M., Kwak, M. K., Dolan, P. M., Itoh, K., Yamamoto, M., Talalay, P. and Kensler, T. W. (2001) Sensitivity to carcinogenesis is increased and chemoprotective efficacy of enzyme inducers is lost in nrf2 transcription factor-deficient mice. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3410-3415
- Iida, K., Itoh, K., Kumagai, Y., Oyasu, R., Hattori, K., Kawai, K., Shimazui, T., Akaza, H. and Yamamoto, M. (2004) Nrf2 is essential for the chemopreventive efficacy of oltipraz against urinary bladder carcinogenesis. *Cancer Res.* **64**, 6424-6431
- McMahon, M., Itoh, K., Yamamoto, M., Chanas, S. A., Henderson, C. J., McLellan, L. I., Wolf, C. R., Cavin, C. and Hayes, J. D. (2001) The Cap'n'Collar basic leucine zipper transcription factor Nrf2 (NF-E2

- p45-related factor 2) controls both constitutive and inducible expression of intestinal detoxification and glutathione biosynthetic enzymes. *Cancer Res.* **61**, 3299-3307
- 16 Luna, L., Skammelsrud, N., Johnsen, O., Abel, K. J., Weber, B. L., Prydz, H. and Kolsto, A. B. (1995) Structural organization and mapping of the human TCF11 gene. *Genomics* **27**, 237-244
  - 17 Luna, L., Johnsen, O., Skartlien, A. H., Pedetour, F., Turc-Carel, C., Prydz, H. and Kolsto, A. B. (1994) Molecular cloning of a putative novel human bZIP transcription factor on chromosome 17q22. *Genomics* **22**, 553-562
  - 18 Chan, J. Y., Han, X. L. and Kan, Y. W. (1993) Cloning of Nrf1, an NF-E2-related transcription factor, by genetic selection in yeast. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 11371-11375
  - 19 Husberg, C., Murphy, P., Martin, E. and Kolsto, A. B. (2001) Two domains of the human bZIP transcription factor TCF11 are necessary for transactivation. *J. Biol. Chem.* **276**, 17641-17652
  - 20 McKie, J., Johnstone, K., Mattei, M. G. and Scambler, P. (1995) Cloning and mapping of murine Nfe211. *Genomics* **25**, 716-719
  - 21 Furukawa, M. and Xiong, Y. (2005) BTB protein Keap1 targets antioxidant transcription factor Nrf2 for ubiquitination by the Cullin 3-Roc1 ligase. *Mol. Cell. Biol.* **25**, 162-171
  - 22 Zhang, D. D., Lo, S. C., Cross, J. V., Templeton, D. J. and Hannink, M. (2004) Keap1 is a redox-regulated substrate adaptor protein for a Cul3-dependent ubiquitin ligase complex. *Mol. Cell. Biol.* **24**, 10941-10953
  - 23 McMahan, M., Itoh, K., Yamamoto, M. and Hayes, J. D. (2003) Keap1-dependent proteasomal degradation of transcription factor Nrf2 contributes to the negative regulation of antioxidant response element-driven gene expression. *J. Biol. Chem.* **278**, 21592-21600
  - 24 Cullinan, S. B., Gordan, J. D., Jin, J., Harper, J. W. and Diehl, J. A. (2004) The Keap1-BTB protein is an adaptor that bridges Nrf2 to a Cul3-based E3 ligase: oxidative stress sensing by a Cul3-Keap1 ligase. *Mol. Cell. Biol.* **24**, 8477-8486
  - 25 Dinkova-Kostova, A. T., Holtzclaw, W. D. and Kensler, T. W. (2005) The role of Keap1 in cellular protective responses. *Chem. Res. Toxicol.* **18**, 1779-1791
  - 26 Itoh, K., Wakabayashi, N., Katoh, Y., Ishii, T., Igarashi, K., Engel, J. D. and Yamamoto, M. (1999) Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain. *Genes Dev.* **13**, 76-86
  - 27 McMahan, M., Thomas, N., Itoh, K., Yamamoto, M. and Hayes, J. D. (2006) Dimerisation of substrate adaptors can facilitate cullin-mediated ubiquitylation of proteins by a 'tethering' mechanism: A two-site interaction model for the Nrf2-Keap1 complex. *J. Biol. Chem.* 2006 Jun 21 (Epub)
  - 28 Katoh, Y., Iida, K., Kang, M. I., Kobayashi, A., Mizukami, M., Tong, K. I., McMahan, M., Hayes, J. D., Itoh, K. and Yamamoto, M. (2005) Evolutionary conserved N-terminal domain of Nrf2 is essential for the Keap1-mediated degradation of the protein by proteasome. *Arch. Biochem. Biophys.* **433**, 342-350
  - 29 Kobayashi, A., Ohta, T. and Yamamoto, M. (2004) Unique function of the Nrf2-Keap1 pathway in the inducible expression of antioxidant and detoxifying enzymes. *Methods Enzymol.* **378**, 273-286
  - 30 Zhang, Y., Cho, Y. Y., Petersen, B. L., Bode, A. M., Zhu, F. and Dong, Z. (2003) Ataxia telangiectasia mutated proteins, MAPKs, and RSK2 are involved in the phosphorylation of STAT3. *J. Biol. Chem.* **278**, 12650-12659
  - 31 Solomin, L., Johansson, C. B., Zetterstrom, R. H., Bissonnette, R. P., Heyman, R. A., Olson, L., Lendahl, U., Frisen, J. and Perlmann, T. (1998) Retinoid-X receptor signalling in the developing spinal cord. *Nature* **395**, 398-402
  - 32 Nioi, P., McMahan, M., Itoh, K., Yamamoto, M. and Hayes, J. D. (2003) Identification of a novel Nrf2-regulated antioxidant response element (ARE) in the mouse NAD(P)H:quinone oxidoreductase 1 gene: reassessment of the ARE consensus sequence. *Biochem. J.* **374**, 337-348
  - 33 Hudson, F. N. and Kavanagh, T. J. (2000) Cloning and characterization of the proximal promoter region of the mouse glutamate-L-cysteine ligase regulatory subunit gene. *Biochim. Biophys. Acta* **1492**, 447-451
  - 34 Mignotte, V., Wall, L., deBoer, E., Grosveld, F. and Romeo, P. H. (1989) Two tissue-specific factors bind the erythroid promoter of the human porphobilinogen deaminase gene. *Nucleic Acids Res.* **17**, 37-54
  - 35 Rushmore, T. H., Morton, M. R. and Pickett, C. B. (1991) The antioxidant responsive element. Activation by oxidative stress and identification of the DNA consensus sequence required for functional activity. *J. Biol. Chem.* **266**, 11632-11639
  - 36 Wakabayashi, N., Itoh, K., Wakabayashi, J., Motohashi, H., Noda, S., Takahashi, S., Imakado, S., Kotsuji, T., Otsuka, F., Roop, D. R., Harada, T., Engel, J. D. and Yamamoto, M. (2003) Keap1-null mutation leads to postnatal lethality due to constitutive Nrf2 activation. *Nat. Genet.* **35**, 238-245

- 37 Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685
- 38 Johnsen, O., Murphy, P., Prydz, H. and Kolsto, A. B. (1998) Interaction of the CNC-bZIP factor TCF11/LCR-F1/Nrf1 with MafG: binding-site selection and regulation of transcription. *Nucleic Acids Res.* **26**, 512-520
- 39 Kozak, M. (2002) Pushing the limits of the scanning mechanism for initiation of translation. *Gene* **299**, 1-34
- 40 Murphy, P. and Kolsto, A. (2000) Expression of the bZIP transcription factor TCF11 and its potential dimerization partners during development. *Mech. Dev.* **97**, 141-148
- 41 Smith, M. J. and Koch, G. L. (1989) Multiple zones in the sequence of calreticulin (CRP55, calregulin, HACBP), a major calcium binding ER/SR protein. *EMBO J.* **8**, 3581-3586
- 42 Brown, M. S. and Goldstein, J. L. (1997) The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. *Cell* **89**, 331-340
- 43 Haze, K., Yoshida, H., Yanagi, H., Yura, T. and Mori, K. (1999) Mammalian transcription factor ATF6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress. *Mol. Biol. Cell.* **10**, 3787-3799
- 44 Stirling, J. and O'Hare, P. (2006) CREB4, a transmembrane bZip transcription factor and potential new substrate for regulation and cleavage by S1P. *Mol. Biol. Cell.* **17**, 413-426
- 45 Zhang, K., Shen, X., Wu, J., Sakaki, K., Saunders, T., Rutkowski, D. T., Back, S. H. and Kaufman, R. J. (2006) Endoplasmic reticulum stress activates cleavage of CREBH to induce a systemic inflammatory response. *Cell* **124**, 587-599
- 46 Ye, J., Rawson, R. B., Komuro, R., Chen, X., Dave, U. P., Prywes, R., Brown, M. S. and Goldstein, J. L. (2000) ER stress induces cleavage of membrane-bound ATF6 by the same proteases that process SREBPs. *Mol. Cell* **6**, 1355-1364
- 47 Cheng, D., Espenshade, P. J., Slaughter, C. A., Jaen, J. C., Brown, M. S. and Goldstein, J. L. (1999) Secreted site-1 protease cleaves peptides corresponding to luminal loop of sterol regulatory element-binding proteins. *J. Biol. Chem.* **274**, 22805-22812
- 48 Toure, B. B., Munzer, J. S., Basak, A., Benjannet, S., Rochemont, J., Lazure, C., Chretien, M. and Seidah, N. G. (2000) Biosynthesis and enzymatic characterization of human SKI-1/S1P and the processing of its inhibitory prosegment. *J. Biol. Chem.* **275**, 2349-2358
- 49 Martoglio, B. and Dobberstein, B. (1998) Signal sequences: more than just greasy peptides. *Trends Cell. Biol.* **8**, 410-415
- 50 Husberg, C., Murphy, P., Bjorgo, E., Kalland, K. H. and Kolsto, A. B. (2003) Cellular localisation and nuclear export of the human bZIP transcription factor TCF11. *Biochim. Biophys. Acta* **1640**, 143-151
- 51 McMahan, M., Thomas, N., Itoh, K., Yamamoto, M. and Hayes, J. D. (2004) Redox-regulated turnover of Nrf2 is determined by at least two separate protein domains, the redox-sensitive Neh2 degron and the redox-insensitive Neh6 degron. *J. Biol. Chem.* **279**, 31556-31567
- 52 Kobayashi, M., Itoh, K., Suzuki, T., Osanai, H., Nishikawa, K., Katoh, Y., Takagi, Y. and Yamamoto, M. (2002) Identification of the interactive interface and phylogenetic conservation of the Nrf2-Keap1 system. *Genes Cells* **7**, 807-820
- 53 Velichkova, M. and Hasson, T. (2005) Keap1 regulates the oxidation-sensitive shuttling of Nrf2 into and out of the nucleus via a Crm1-dependent nuclear export mechanism. *Mol. Cell. Biol.* **25**, 4501-4513
- 54 Motohashi, H. and Yamamoto, M. (2004) Nrf2-Keap1 defines a physiologically important stress response mechanism. *Trends Mol. Med.* **10**, 549-557
- 55 Padmanabhan, B., Tong, K. I., Ohta, T., Nakamura, Y., Scharlock, M., Ohtsuji, M., Kang, M. I., Kobayashi, A., Yokoyama, S. and Yamamoto, M. (2006) Structural basis for defects of keap1 activity provoked by its point mutations in lung cancer. *Mol. Cell* **21**, 689-700
- 56 Zhang, D.D., Lo, S.-C., Cross, J.V., Templeton, D.J. and Hannink, M. (2004) Keap1 is a redox-regulated substrate adaptor protein for a Cul3-dependent ubiquitin ligase complex. *Mol. Cell. Biol.* **24**, 10941-10953
- 57 Tong, K. I., Katoh, Y., Kusunoki, H., Itoh, K., Tanaka, T. and Yamamoto, M. (2006) Keap1 recruits Neh2 through binding to ETGE and DLG motifs: characterization of the two-site molecular recognition model. *Mol. Cell. Biol.* **26**, 2887-2900

## FOOTNOTES

¶ To whom correspondence should be addressed: Dr. Yiguo Zhang and Prof. John D. Hayes, Biomedical Research Centre, Ninewells Hospital and Medical School, University of Dundee, Dundee DD1 9SY, Scotland, United Kingdom. Tel: +44 1382 632788; Fax: +44 1382 669993 and E-mail: [y.z.zhang@dundee.ac.uk](mailto:y.z.zhang@dundee.ac.uk) or [john.hayes@cancer.org.uk](mailto:john.hayes@cancer.org.uk).

<sup>1</sup>The abbreviations used are: ARE, antioxidant response element; bZIP, basic region-leucine zipper;  $\beta$ -gal, galactosidase; CNC, cap 'n' collar; CMV, cytomegalovirus; DAPI, 4',6-Diamidino-2-phenylindole; DIC, differential interference contrast transmission; FITC, fluorescein isothiocyanate; Gal4D, Gal4 DNA-binding domain; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; GCLM, glutamate cysteine ligase modifier; Keap1, Kelch-like ECH-associated protein 1; KO, knockout; LCR-F1, locus control region-factor 1; LDS, lithium dodecyl-sulphate; Luc, luciferase; MEF, mouse embryonic fibroblast; Neh2L, Nrf2-ECH homology 2-like; NQO1, NAD(P)H:quinone oxidoreductase 1; Nrf1, Nrf2, and Nrf3, Nuclear Factor-Erythroid 2 p45 subunit-related factors 1, 2 and 3; NTD, N-terminal domain; NuPAGE, Novex® polyacrylamide gel electrophoresis; *P*, promoter; PBGD, porphobilinogen deaminase; SV40, simian virus 40; TCF11, transcription factor 11; TK, thymidine kinase promoter; UAS, a Gal4-specific binding site.

## FIGURE LEGENDS

### Figure 1 Nrf1 is less potent than Nrf2 at transactivating ARE-driven reporter gene expression

(**A** and **B**) The core ARE in various luciferase gene promoter regions is indicated by a horizontal bar above the sequence. (**A**) The mutated nucleotides around the ARE (mut1) of *P*<sub>-1016/nqo1</sub>-ARE-Luc are underlined. COS-1 cells were cotransfected with 0.6  $\mu$ g of each reporter construct for either the wild-type *P*<sub>-1016/nqo1</sub>-ARE-Luc or the mut1 variant together with 1.1  $\mu$ g of pcDNA3.1/V5His B constructs encoding Nrf1 or Nrf2; controls included an empty pcDNA3.1/V5His B vector, and 0.3  $\mu$ g of pRL-TK Renilla. (**B**) Further similar experiments were performed in COS-1 cells with the 41-bp reporter constructs *P*<sub>SV40/nqo1</sub>-ARE-Luc, *P*<sub>SV40/gclm</sub>-ARE-Luc or *P*<sub>SV40/pbgd</sub>-ARE-Luc, together with 1.1  $\mu$ g of expression constructs for Nrf1, Nrf2 or the empty pcDNA3.1/V5His B vector, and 0.3  $\mu$ g of pRL-TK Renilla. In both panels **A** and **B**, luciferase reporter gene activity was measured in cell lysates about 36 h after transfection. Empty pcDNA3 and pGL3-Promoter luciferase vectors (corrected for transfection efficiency by Renilla activity) were used as controls, and the luciferase value obtained was subtracted from experimental results. Transactivation by Nrf1 or Nrf2 is represented as a fold change (mean  $\pm$  S.D.) from three independent experiments. (**C**) The proteins in the above COS-1 cell lysates were resolved by either discontinuous SDS-PAGE, containing 8% polyacrylamide in the upper half of the resolving gel and 12% polyacrylamide in the lower half of the gel (*left panel*) or continuous SDS-PAGE in 8% polyacrylamide gels (*right upper panel*). Immunoblotting was performed with antibodies against V5, or GAPDH; the latter served as an internal protein loading control. Specific cross-reacting bands for

mouse Nrf1 or mouse Nrf2 are marked (\*). The mobility of pre-stained protein standards is shown to allow estimation of molecular mass (kDa).

### Figure 2 Nrf1 is negatively controlled by its NTD

(A) Expression constructs for wild-type Nrf1, Nrf1<sup>Δ2-120</sup>, Nrf1<sup>Δ2-150</sup>, or Nrf1<sup>Δ2-170</sup>, each along with *P<sub>SV40</sub>nqo1*-ARE-Luc and β-gal reporters were transfected into COS-1. Luciferase activity was measured and normalized to that of β-gal. Data are shown as a fold change (mean ± S.D.) calculated from duplicates in three independent experiments. Ectopic expression of C-terminally V5-tagged protein in COS-1 cells was examined by western blotting following resolution by 7% LDS-NuPAGE (right panel). (B) The Gal4 DNA-binding (Gal4D) fusion proteins with full-length Nrf1<sup>1-741</sup>, its C-terminal CNC plus the bZIP and CTD regions (i.e. Nrf1<sup>607-741</sup>), or different N-terminal portions based on Nrf1<sup>1-607</sup>, were generated and are schematically represented in the *left upper panels*. The constructs encoding various Gal4D/Nrf1-V5 fusion proteins were cotransfected into COS-1 cells along with a Gal4-specific UAS (4×)-Luc reporter and pRL-TK Renilla. Approximately 36 h after transfection, reporter gene activity was measured. The fold activation of UAS-driven luciferase activity was calculated from duplicates and is shown as the mean value ± S.D. from three independent experiments. Proteins expressed from these Gal4D/Nrf1 fusion constructs were resolved by discontinuous SDS-PAGE.

### Figure 3 The ETGE and DIDLID/DLG motifs in Nrf1 neither regulate its activity nor control its subcellular localization

(A) Diagrammatic representation of constructs expressing wild-type Nrf1 and deletion mutants lacking sequences around the Neh2L subdomain. (B) These Nrf1 expression constructs, together with pTK-ARE-Luc or *P<sub>SV40</sub>nqo1*-ARE-Luc, and β-gal reporters, were cotransfected into either COS-1 or RL-34 cells. Thirty-six hours after transfection, luciferase reporter gene activity was measured. The results represent fold changes (mean ± S.D.) from the mean activity measured in three separate wells from three independent experiments. Ectopic expression of proteins in the above cell lysates was determined using 7% LDS-NuPAGE followed by immunoblotting with a V5 antibody (lower panels). Samples loaded in each well contained equal amounts of β-gal activity. (C) Immunocytochemistry was performed in COS-1 cells that were transfected with 1.3 μg of expression constructs for wild-type Nrf1 or the mutants Nrf1<sup>ΔETGE</sup> and Nrf1<sup>ΔDIDLID/DLG</sup>, as well as Nrf2/pcDNA3.1/V5His B or empty vector. Confocal images were obtained using a LSM 510 laser scanning microscope. FITC (fluorescein isothiocyanate) labelled second antibody was used to locate V5-tagged Nrf1 or Nrf2. The DAPI signal represents nuclear DNA. DIC signifies images obtained by normal light microscopy. The merge signal represents the results obtained when the three images were superimposed. The length of the horizontal bar = 20 μm. (D) The cytoplasmic-nuclear distribution of the FITC signals was quantified by calculating the percentage of cells (100 cells were counted) with predominantly cytoplasmic (C) and nuclear (N) staining.

### Figure 4 Keap1 does not negatively regulate Nrf1

(A) Three different experimental strategies were employed to determine whether Keap1 regulates the ability of Nrf1 to transactivate ARE-driven gene expression. Firstly (A, left panel), COS-1 cells were transfected, using Lipofectin reagent (5  $\mu$ l) in 1 ml of Opti-MEM1, with 0.6  $\mu$ g of an expression vector for either Keap1 or empty pcDNA3.1/V5His B, along with 0.6  $\mu$ g of an expression construct for either Nrf1 or Nrf2 or an empty pcDNA3.1/V5His B vector, and 0.6  $\mu$ g of  $P_{SV40}nqo1$ -ARE-Luc and 0.2  $\mu$ g of pRL-TK Renilla. Secondly (A, middle panel), *keap1*<sup>+/+</sup> and *keap1*<sup>-/-</sup> MEFs were transfected, using 10  $\mu$ l of Lipofectamine 2000 and 2.5 ml of Opti-MEM1, with 2.6  $\mu$ g of an expression construct for either Nrf2, Nrf1, Nrf1 <sup>$\Delta$ 2-120</sup> or the empty pcDNA3.1/V5His B vector, together with 1.2  $\mu$ g of pTK-ARE-Luc and 0.2  $\mu$ g of  $\beta$ -gal reporter. Thirdly (A, right panel), *keap1*<sup>-/-</sup> MEFs were transfected, using Lipofectamine 2000, with a total of 3.8  $\mu$ g of DNA comprising 1.3  $\mu$ g of either Keap1 or empty pcDNA3.1/V5His B, along with 1.3  $\mu$ g of an expression construct for either Nrf2, Nrf1, Nrf1 <sup>$\Delta$ 2-120</sup>, or empty pcDNA3.1/V5His B, and 1.0  $\mu$ g of pGL-6 $\times$ ARE-Luc and 0.2  $\mu$ g of  $\beta$ -gal reporter. Luciferase activity was measured and normalized to activity of Renilla (*left panel*) or  $\beta$ -gal (*middle and right panels*). It is presented as a fold change (mean  $\pm$  S.D.) from two independent experiments each performed in triplicate. (B) The *keap1*<sup>+/+</sup> and *keap1*<sup>-/-</sup> MEFs were transfected with 2  $\mu$ g of an expression construct either for wild-type Nrf1, its mutants Nrf1 <sup>$\Delta$ ETGE</sup> or Nrf1 <sup>$\Delta$ DIDLID/DLG</sup>, or the empty pcDNA3.1/V5His B plasmid. These V5-tagged Nrf1 proteins were visualized using immunocytochemistry with a FITC-coated secondary antibody. Subsequently, related images were analyzed by confocal microscopy as described in the legend for Figure 3. The horizontal bar represents 20  $\mu$ m.

### Figure 5 The NTD of Nrf1 is required for its association with the endoplasmic reticulum

COS-1 (A) and RL-34 cells (B) were transfected with expression constructs (1.3  $\mu$ g) for either wild-type Nrf1 or its Nrf1 <sup>$\Delta$ 2-120</sup> mutant alone or together with pDsRed2-ER encoding the fusion protein ER/DsRed. Confocal imaging was performed followed by immunocytochemistry with a primary antibody against V5 and a FITC-conjugated secondary antibody (*green image panels* in A and B). Red images that represented ER/DsRed were also measured (excitation 558 nm, emission 583 nm) (B). Other images for DAPI, DIC and Merge were obtained as described in the legend for Figure 3. A ratio of the subcellular distribution of the various ectopically expressed Nrf1 proteins is represented as a percentage of positively stained cells (*lower panels* in A and B) that gave cytoplasmic (C) and nuclear (N) staining. Bar = 20  $\mu$ m. Localization of V5-tagged Nrf2 was used as a control.

### Figure 6 Attachment of the NTD from Nrf1 to Nrf2 results in negative regulation of the resulting fusion proteins and their targeting to the endoplasmic reticulum

(A) *Keap1*<sup>-/-</sup> MEFs were cotransfected with 2  $\mu$ g of a pcDNA3.1/V5His B expression construct for either Nrf1, Nrf2, NTD/Nrf2<sup>f</sup> or NTD/Nrf2<sup>s</sup>, together with 1  $\mu$ g of pDsRed2-ER. Confocal imaging was performed as described in the legends of Figures 3 and 5. Bar = 20  $\mu$ m. (B) Shows *keap1*<sup>+/+</sup> and *keap1*<sup>-/-</sup> MEFs that were cotransfected with 2.6  $\mu$ g of an expression construct for either Nrf2, NTD/Nrf2<sup>f</sup>, NTD/Nrf2<sup>s</sup>, or empty

pcDNA3.1/V5His B, along with 1.2  $\mu\text{g}$  of pTK-ARE-Luc and 0.2  $\mu\text{g}$  of  $\beta$ -gal reporter plasmids. **(C)** Shows *keap1*<sup>-/-</sup> MEFs cotransfected with 3.8  $\mu\text{g}$  of DNA comprising 1.3  $\mu\text{g}$  of either Keap1 expression construct or empty pcDNA3.1/V5His B, together with 1.3  $\mu\text{g}$  of an expression construct for either Nrf2, NTD/Nrf2<sup>f</sup>, Nrf1 <sup>$\Delta$ 2-120</sup> or empty pcDNA3.1/V5His B, 1.0  $\mu\text{g}$  of pGL-6 $\times$ ARE-Luc and 0.2  $\mu\text{g}$  of  $\beta$ -gal reporter plasmids. Luciferase activity was measured and normalized to  $\beta$ -gal activity, and is presented as a fold change (mean  $\pm$  S.D.) from two independent experiments performed in triplicate. **(D)** Cotransfection into COS-1 cells with each of the indicated Gal4D/Nrf1 fusion constructs with pDsRed2-ER, immunocytochemistry and confocal imaging were performed as described in the legend to Figure 3.

## Table 1

**Table 1 Protein domains in mouse Nrf1**

Domain	Residues	Salient features*
NTD, N-terminal domain	1-124	contains several discrete hydrophobic peptide sequences
AD1, acidic domain 1	125-295	equivalent to "N-terminal acidic-rich region" in TCF11; contains subdomains resembling Neh2 (i.e. residues 156-242) and Neh5 (i.e. residues 275-295) in Nrf2
NST, Asn/Ser/Thr-rich	296-403	contains 27% S+T, 14% Q+N and 6% D+E
AD2, acidic domain 2	404-453	equivalent to "central acidic-rich region" in TCF11 and the acidic region in LCR-F1
SR, serine-repeat	454-488	contains 77% S; equivalent to the "serine-rich" regions of TCF11 and LCR-F1
Neh6L, Neh6-like	489-580	contains 19% S+P, 17% D+E and 7% K+R; it resembles Neh6 in Nrf2
CNC, cap 'n' collar	581-624	shares 64% identity with the same domain in Nrf2
bZIP, basic-region zipper	625-686	shares 56% identity with the same domain in Nrf2
CTD, C-terminal domain	687-741	shares 53% identity with the Neh3 domain in Nrf2

\*Single capital letters in the column represent amino acids.

Fig 1

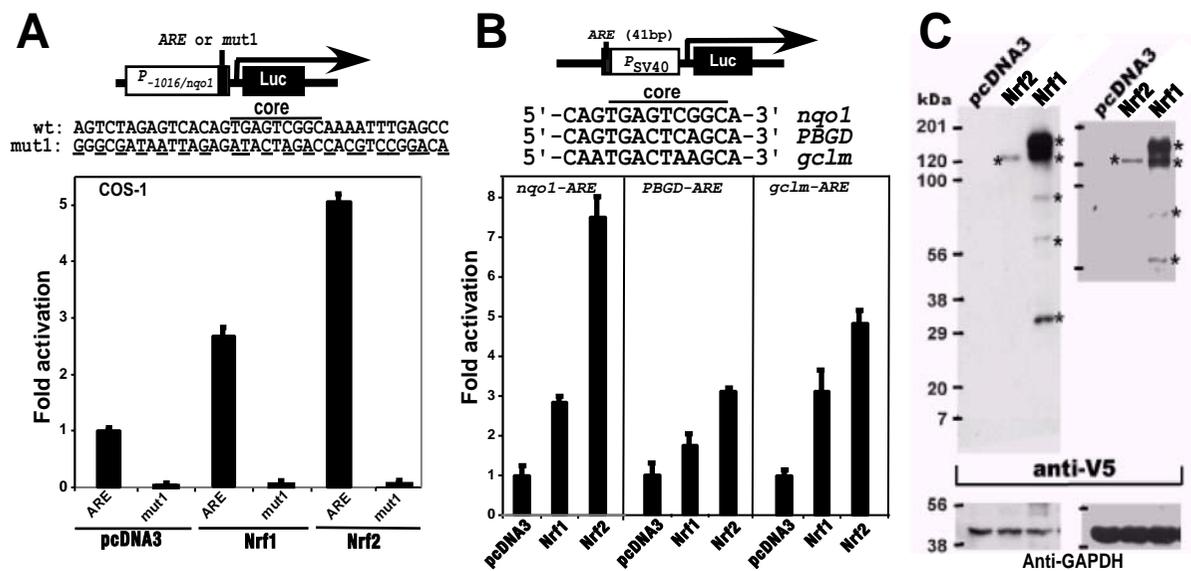


Fig 2

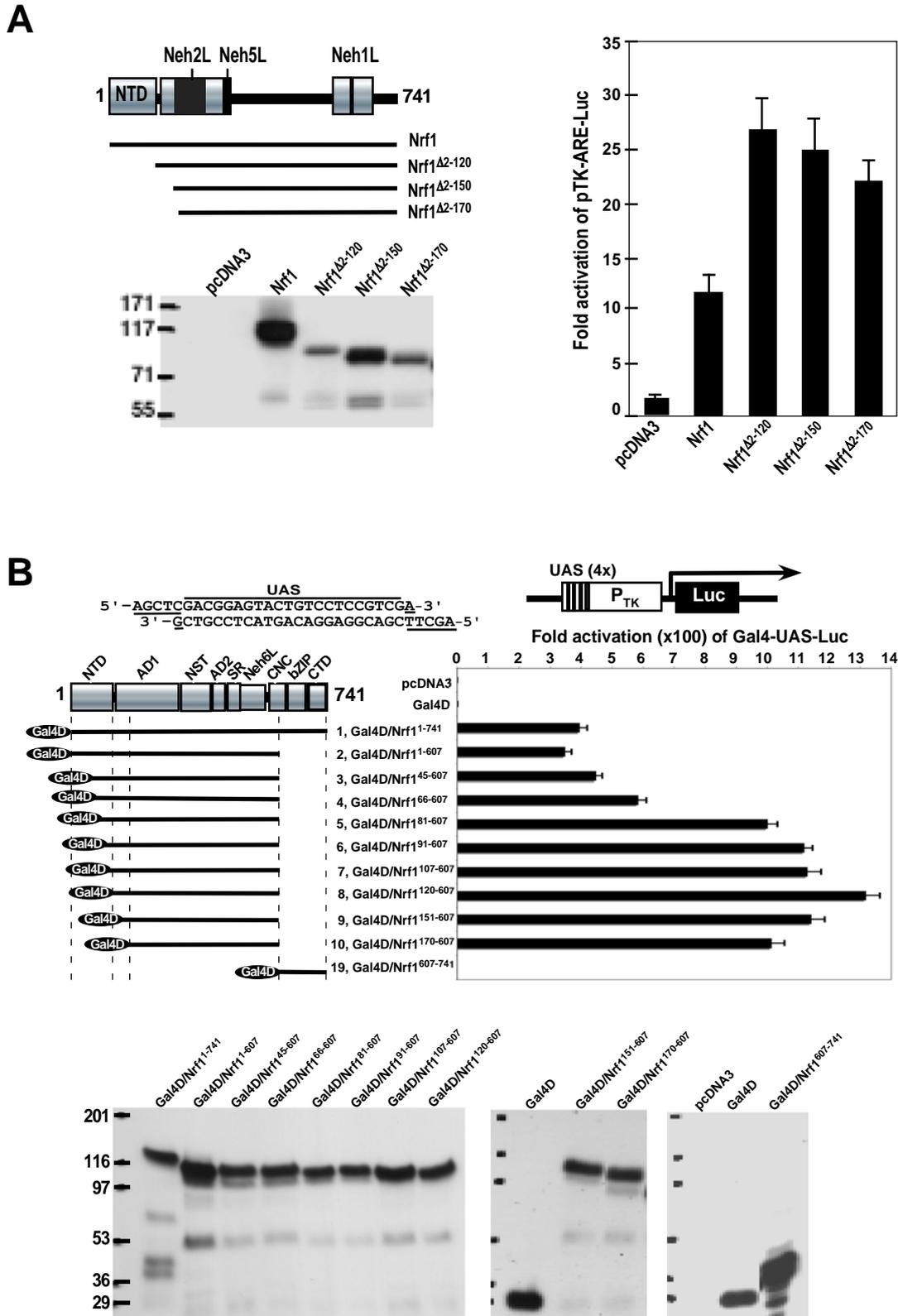


Fig 3

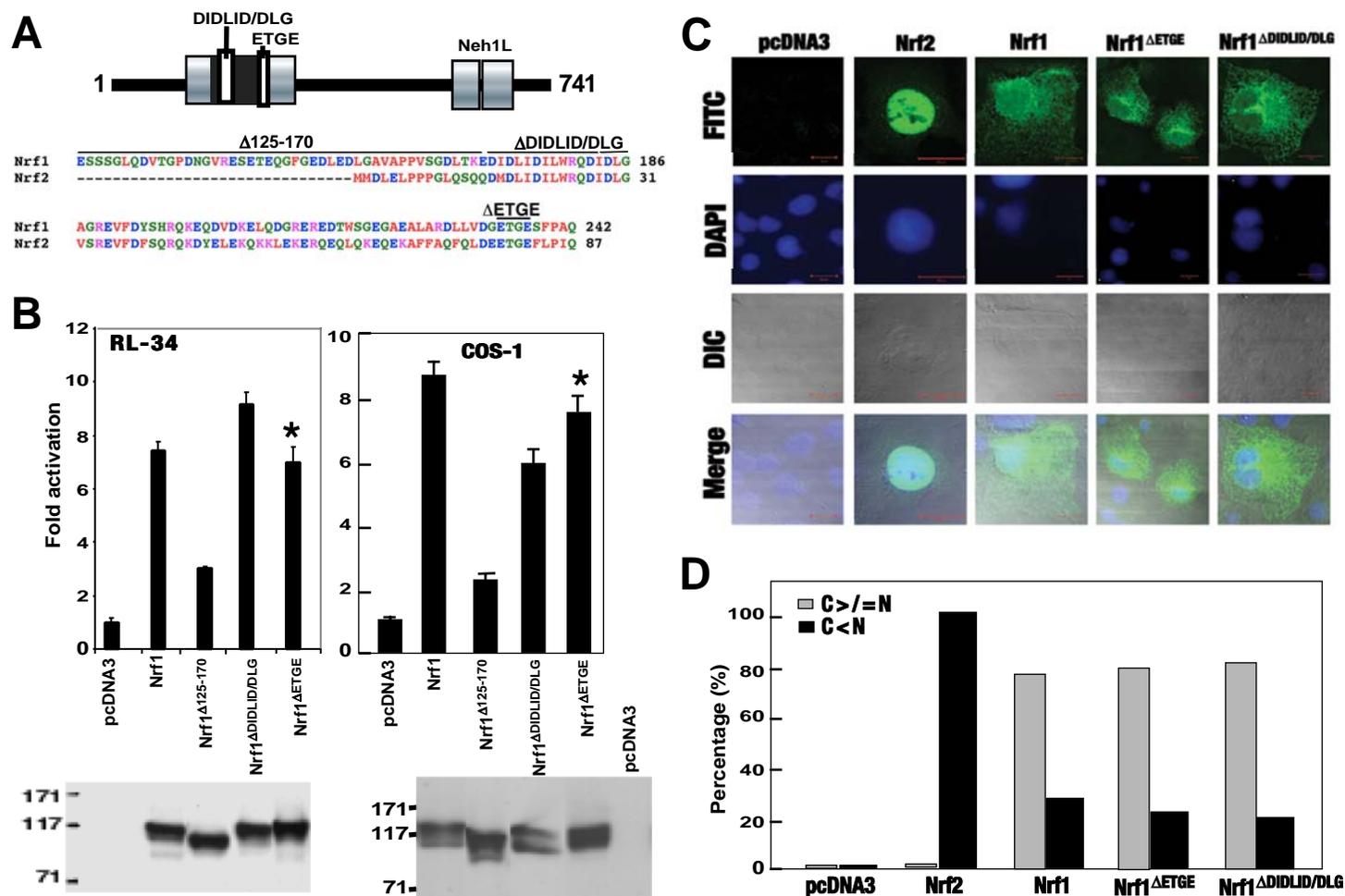


Fig 4

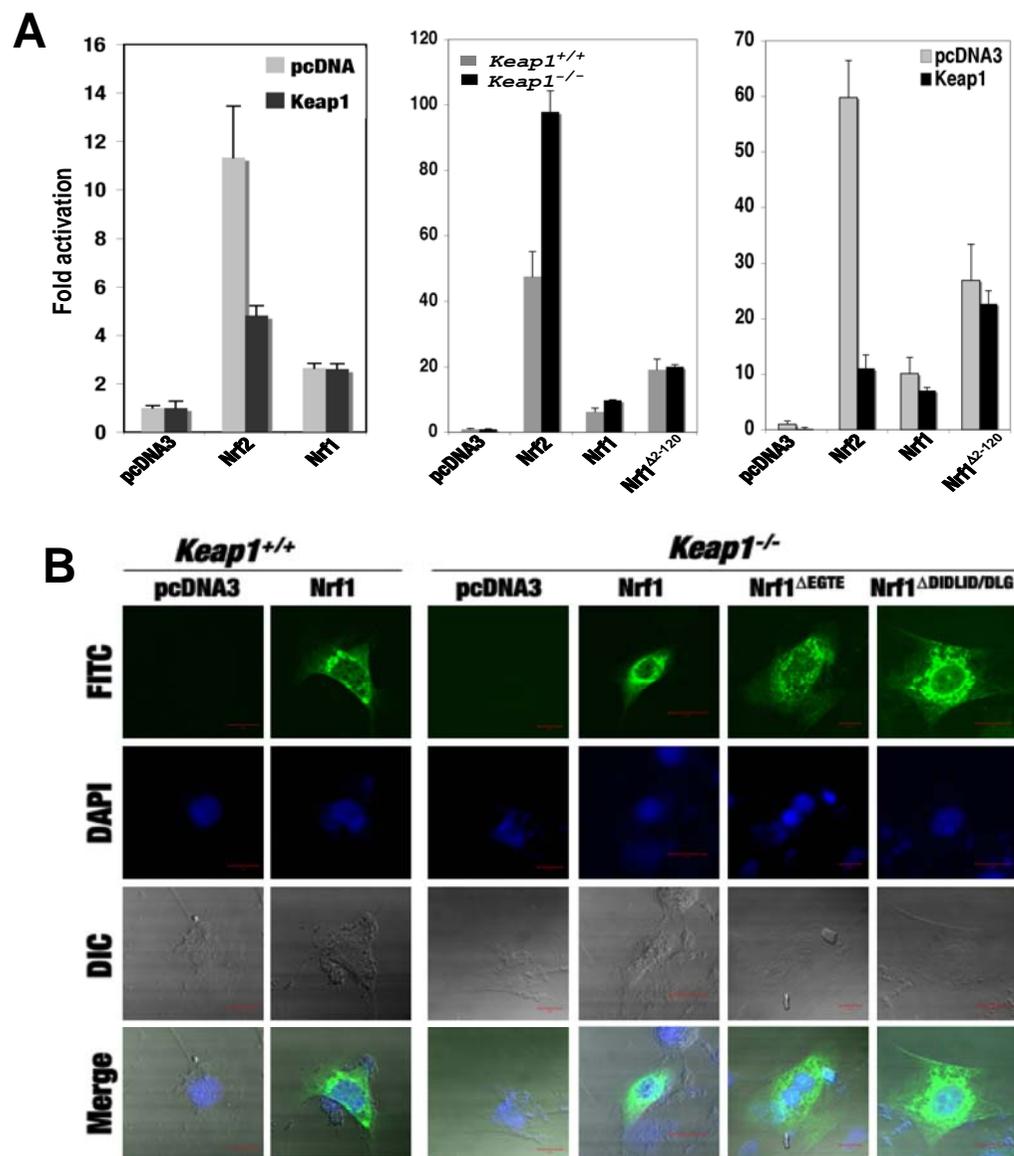


Fig 5

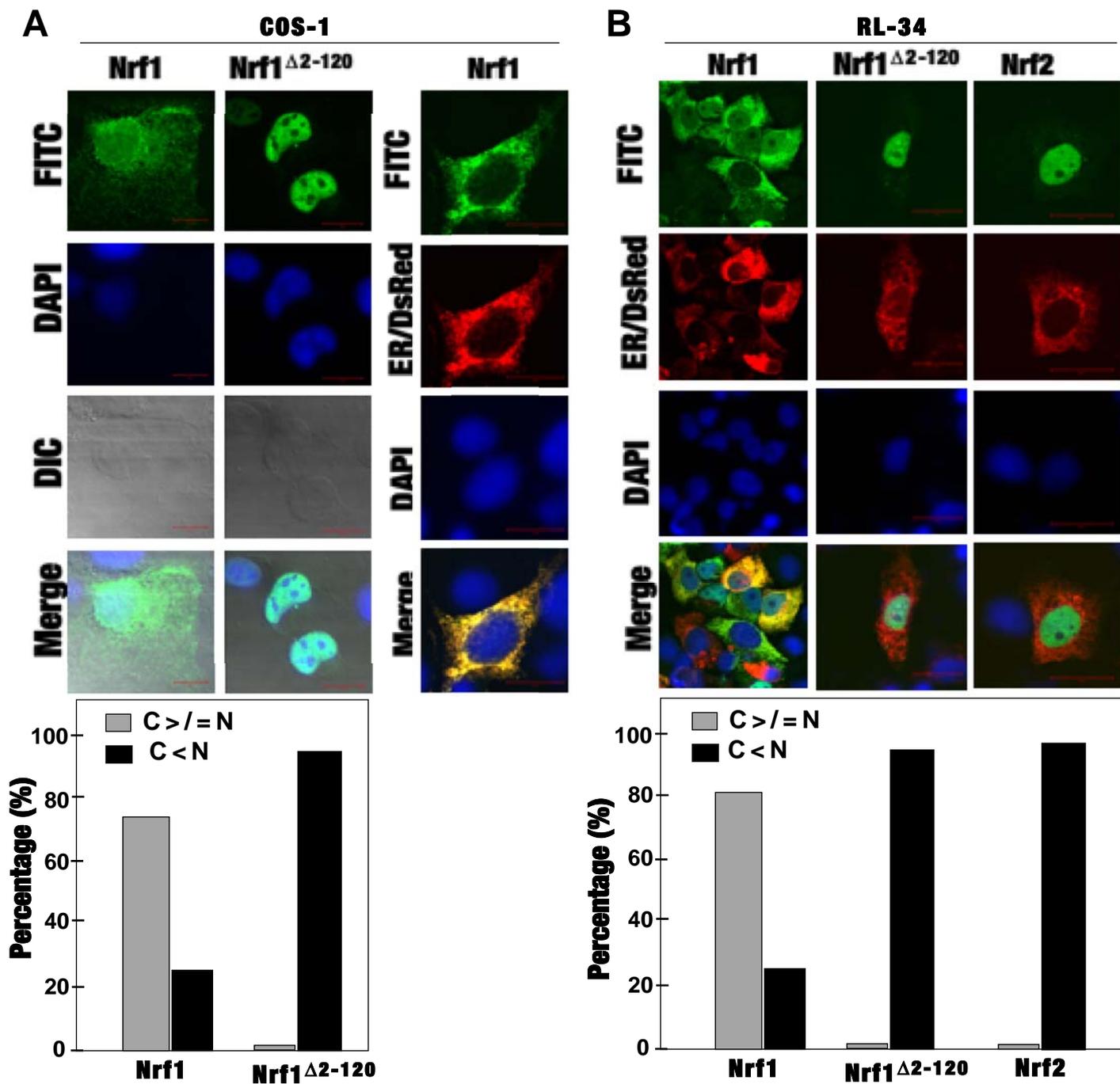


Fig 6

